



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **07295595 A**(43) Date of publication of application: **10 . 11 . 95**(51) Int. Cl. **G10L 9/14**  
**G10L 9/18**(21) Application number: **06088206**(22) Date of filing: **26 . 04 . 94**(71) Applicant: **MATSUSHITA ELECTRIC IND CO LTD**(72) Inventor: **TANAKA NAOYA**(54) **SPEECH ENCODING METHOD**

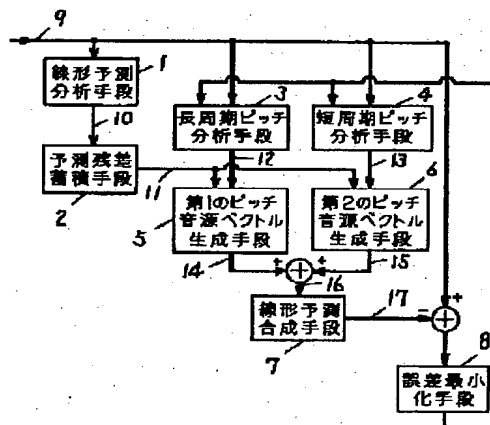
## (57) Abstract:

**PURPOSE:** To provide the speech encoding method which improves the power of expression of the pitch sound source vector of an input speech having a different pitch period without increasing the amount of information.

**CONSTITUTION:** The input speech 9 is analyzed by a long-period pitch analyzing means 3 and a short-period pitch analyzing means 4 to detect a long-period pitch 12 and a short-period pitch 13, and prediction residues matching the pitches are selected from a prediction residue storage means 2 to generate a long-period pitch sound source vector 14 and a short-period pitch sound source vector 15, which are put together to generate a pitch sound source vector 16. A synthesized speech 17 is generated by using the pitch sound source vector 16 and fed back to the pitch analyzing means so that the difference from the input speech 9 is minimized by an error minimizing means 8; and a pitch sound source vector corresponding to the input speech 9 having the different pitch period is generated and the frequency of the long-period pitch analysis and the frequency of the short-period pitch analysis are set to mutually different values in a speech frame, so that the

information quantity is prevented from increasing.

COPYRIGHT: (C)1995,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-295595

(43) 公開日 平成7年(1995)11月10日

(51) Int.Cl.<sup>6</sup>

G 1 0 L 9/14

9/18

識別記号

G

J

E

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数 2 O L (全 4 頁)

(21) 出願番号 特願平6-88206

(22) 出願日 平成6年(1994)4月26日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 田中 直也

神奈川県横浜市港北区綱島東四丁目3番1

号 松下通信工業株式会社内

(74) 代理人 弁理士 栗野 重孝

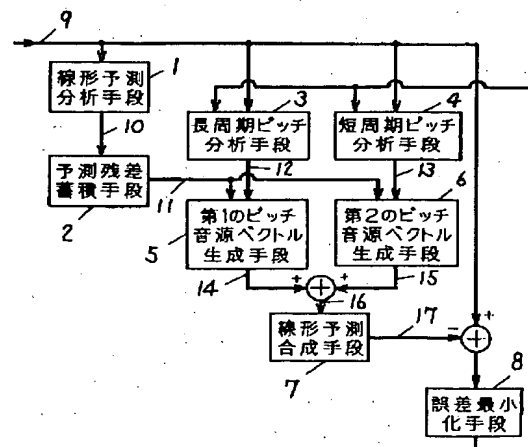
(54) 【発明の名称】 音声符号化方法

(57) 【要約】

【目的】 異なるピッチ周期を有する入力音声のピッチ音源ベクトルの表現力を、情報量を増加させずに向上させた音声符号化方法を提供する。

【構成】 入力音声9を長周期ピッチ分析手段3と短周期ピッチ分析手段4により分析して長周期ピッチ12と短周期ピッチ13を検出し、それぞれにピッチに合った予測残差を予測残差蓄積手段2から選択して長周期ピッチ音源ベクトル14と短周期ピッチ音源ベクトル15とを生成し、それらの合成によりピッチ音源ベクトル16を生成する。このピッチ音源ベクトル16を用いて合成音声17を生成し、入力音声9との誤差を誤差最小化手段8により最小にするようにピッチ分析手段にフィードバックし、異なるピッチ周期を有する入力音声9に対応したピッチ音源ベクトルを生成し、長周期ピッチ分析の回数と短周期ピッチ周期分析の回数を音声フレーム内で相異なる回数に設定して情報量を増加させないようにする。

- 3 長周期ピッチ分析手段
- 4 短周期ピッチ分析手段
- 9 入力音声
- 10 予測残差
- 11 遅れ予測残差
- 12 長周期ピッチ
- 13 短周期ピッチ
- 14 長周期ピッチ音源ベクトル
- 15 短周期ピッチ音源ベクトル
- 16 ピッチ音源ベクトル
- 17 合成音声



## 【特許請求の範囲】

【請求項 1】 入力音声のピッチ周期を所定の閾値に対して長い周期と短い周期に分けて分析し、前記分析により得た長周期ピッチに対応する長周期ピッチ音源ベクトルと短周期ピッチに対応する短周期ピッチ音源ベクトルとを用いてピッチ音源ベクトルを生成し、かつ、前記ピッチ分析において音声フレーム内での長いピッチ周期の分析の回数と短いピッチ周期の分析の回数とを相異なる回数に設定して符号化するようにした音声符号化方法。

【請求項 2】 長周期ピッチの分析を音声フレームごとに行い、短周期ピッチの分析を前記音声フレームを分割したサブフレームごとに行うようにした請求項 1 記載の音声符号化方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、低ビットレートの音声符号化方式において、音声の 2 つのピッチ周期を効率的に符号化することにより、音声品質を向上させる方法に関する。

## 【0002】

【従来の技術】近年、音声の蓄積や伝送に効率のよい低ビットレートのデジタル符号化が用いられるが、その音声品質が課題である。低ビットレートの音声符号化方法では、入力音声をもとに所定時間長の音声フレームに分割し、各音声フレームに対して線形予測分析を行い、得られた線形予測係数と予測残差を符号化する方法が主流である。この予測残差は音源符号帳に蓄えられた音源ベクトルによって符号化されるが、音源ベクトルを生成する手段の 1 つとして前フレーム以前の予測残差を蓄えておき、それを音声のピッチ周期に合わせて切り出して用いる手段がある（以下、切り出された音源ベクトルをピッチ音源ベクトルと称す）。

【0003】以下、従来のピッチ音源ベクトル生成手段について説明する。図 3 は従来のピッチ音源ベクトル生成手段の構成を示すブロック図である。図において、301 は入力音声 307 を分割した所定時間長の音声フレームについて線形予測分析し、線形予測係数と予測残差 308 とを算出する線形予測分析手段、302 は予測残差 309 を蓄積する予測残差蓄積手段、303 は前記音声フレームにおけるピッチ周期を分析してピッチ周期 310 を抽出するピッチ周期分析手段、304 はピッチ周期 310 に基づいて予測残差蓄積手段 302 に蓄積した予測残差を切り出してピッチ音源ベクトル 311 を生成する音源ベクトル生成手段、305 は前記線形予測係数とピッチ音源ベクトル 311 とにより合成音声を 312 を生成する線形予測合成手段、306 は合成音声 312 を入力音声 307 と比較検出した誤差をピッチ周期分析手段にフィードバックしてその誤差が最小になるようにする誤差最小化手段である。

【0004】上記構成において、その動作を説明する。

入力音声 307 を線形予測分析手段 301 によって線形予測分析し、その予測残差 308 を予測残差蓄積手段 302 に蓄積する。一方、入力音声 307 をピッチ周期分析手段 303 によって分析し、ピッチ周期 310 を選択する。音源ベクトル生成手段 304 は遅延された予測残差 309 をピッチ周期 310 に合わせて予測残差蓄積手段 302 から切り出すことにより、ピッチ音源ベクトル 311 を生成する。線形予測合成手段 305 はピッチ音源ベクトル 311 を音源として線形予測合成を行い、合成音声を 312 を生成する。誤差最小化手段 306 は入力音声 307 と合成音声を 312 の誤差を最小にするように、ピッチ周期分析手段 303 にフィードバックをかける。以上の過程により最適なピッチ音源ベクトルが生成される。ピッチ周期の分析は音声フレームまたは音声フレームをさらに分割したサブフレームに対して行い、最適なピッチ周期を 1 つ選択する。

## 【0005】

【発明が解決しようとする課題】このような従来の方法では、音声フレームまたはサブフレームに対するピッチ周期が 1 つに限定されるため、複数の話者が同時に話す場合などの、入力音声に複数のピッチ成分が存在する場合には、予測残差を近似するのに最適なピッチ音源ベクトルが得られないという問題があった。

【0006】本発明は上記の課題を解決するもので、2 つのピッチ周期を効率的に符号化し、それを用いてピッチ音源ベクトルを生成することにより音声品質を向上させることを目的とする。

## 【0007】

【課題を解決するための手段】本発明は上記の目的を達成するために、入力音声のピッチ周期を所定の閾値に対して長い周期と短い周期に分けて分析し、前記分析により得た長周期ピッチに対応する長周期ピッチ音源ベクトルと短周期ピッチに対応する短周期ピッチ音源ベクトルとを用いてピッチ音源ベクトルを生成し、かつ、前記ピッチ分析において音声フレーム内での長いピッチ周期の分析の回数と短いピッチ周期の分析の回数とを相異なる回数に設定して符号化するようにした音声符号化方法である。

## 【0008】

【作用】本発明は上記の構成において、入力音声の長周期ピッチに対応する長周期ピッチ音源ベクトルと短周期ピッチに対応する短周期ピッチ音源ベクトルとの合成によるピッチ音源ベクトルが、異なるピッチ周期を有する入力音声に対応するピッチ音源ベクトルとして表現力を向上させ、かつ、長周期ピッチ分析と短周期ピッチ分析の回数を相異なる回数に設定することが 2 種類のピッチ周期を与えるピッチ周期情報の情報量を低減する。

## 【0009】

【実施例】以下、本発明の音声符号化方法の一実施例について図面を参照しながら説明する。図 1 は本実施例の

音声符号化方法の構成を示すブロック図を示す。図において、1は音声入力を分割した所定時間長の音声フレームについて線形予測分析し、線形予測係数と予測残差10とを算出する線形予測分析手段、2は予測残差10を蓄積する予測残差蓄積手段、3は所定の閾値に基づいて前記音声フレームにおける長周期ピッチ12を抽出する長周期ピッチ分析手段、4は前記閾値に基づいて短周期ピッチ13を抽出する短周期ピッチ分析手段、5は長周期ピッチ12に基づいて予測残差蓄積手段2に蓄積した予測残差を切り出して長周期ピッチ音源ベクトル14を生成する第1のピッチ音源ベクトル生成手段、6は同様に短周期ピッチ13に基づいて短周期ピッチ音源ベクトル15を生成する第2のピッチ音源ベクトル生成手段、7は長周期ピッチ音源ベクトル14と短周期音源ベクトル15とを合成したピッチ音源ベクトル16と前記線形予測係数とにより合成音声17を生成する線形予測合成手段、8は合成音声17を入力音声9と比較した誤差を長周期ピッチ分析手段3および短周期ピッチ分析手段4にフィードバックして誤差を最小にする誤差最小化手段、11は遅れ予測残差である。また、図2(a)は入力音声を示す波形図であり、図2(b)は入力音声に対応するピッチ分析のフレームを示す模式図であって、a、b、c、・・・は音声フレーム、1、2、3、・・・は前記音声フレームをさらに分割したサブフレームを示す。

【0010】上記構成において、ピッチ音源ベクトルを生成する動作について説明する。図1において、入力音声9は線形予測分析手段1により所定長の音声フレームについて線形予測分析され、その予測残差10が予測残差蓄積手段2に蓄積される。一方、入力音声9を長周期ピッチ分析手段3および短周期ピッチ分析手段4によりピッチ周期を分析し、それぞれ長周期ピッチ12と短周期ピッチ13を抽出する。第1のピッチ音源ベクトル生成手段5は長周期ピッチ12に基づいて予測残差蓄積手段2に蓄積された予測残差を切り出して長周期ピッチ音源ベクトル14を生成する。また、第2のピッチ音源ベクトル生成手段6は同様に短周期ピッチ13に基づいて短周期ピッチ音源ベクトル15を生成する。これら短周期ピッチ音源ベクトル15と長周期ピッチ音源ベクトル14とを合成してピッチ音源ベクトル16とする。

【0011】線形予測合成手段7はピッチ音源ベクトル16を音源として前記線形予測係数により線形予測合成を行い、合成音声17を生成する。誤差最小化手段8は合成音声17を入力音声9と比較した誤差を、長周期ピッチ分析手段3および短周期ピッチ分析手段4にフィードバックしてその誤差が最小になるようにする。以上の過程により最適な長周期ピッチ12と短周期ピッチ13とが選択され、最適なピッチ音源ベクトル16が合成される。上記動作において、選択された2つのピッチ周期をそのまま符号化するのでは情報量が2倍となってしまう

うので、図2に示したように、長周期ピッチ分析と短周期ピッチ分析の回数を異なるように設定することにより情報量を削減する。すなわち、入力音声101を102のa、b、c、・・・で示される音声フレームと、1、2、3、4で示されるサブフレームとに分割し、長周期ピッチ分析は各音声フレームに対して1回、短周期ピッチ分析は各サブフレームに対して1回行うようにする。ピッチ音源ベクトル16の生成は各サブフレームごとに行うが、長周期ピッチは同一フレーム内のサブフレームに対して同じ値を用い、短周期ピッチは各サブフレームで分析した値を用いる。

【0012】以上のように本実施例の音声符号化方法によれば、入力音声の長周期ピッチと短周期ピッチとを分析し、長周期ピッチ音源ベクトルと短周期ピッチ音源ベクトルとを生成して合成することにより音源ベクトルを生成するとともに、長周期分析と短周期分析の回数を異なる回数に設定することにより、ピッチ周期情報の情報量増加を少なくしてピッチ音源ベクトルの表現力を向上させることができる。

#### 【0013】

【発明の効果】以上の説明から明らかなように、本発明は、入力音声のピッチ周期を所定の閾値に対して長い周期と短い周期に分けて分析し、前記分析により得た長周期ピッチに対応する長周期ピッチ音源ベクトルと短周期ピッチに対応する短周期ピッチ音源ベクトルとを用いてピッチ音源ベクトルを生成し、かつ、前記ピッチ分析において音声フレーム内での長いピッチ周期の分析の回数と短いピッチ周期の分析の回数とを相異なる回数に設定して符号化するようにしたことにより、長周期ピッチと短周期ピッチとに対応したピッチ音源ベクトルを伝送情報量をほとんど増加させずに生成し、音声品質を向上させた音声符号化を実現できる。

#### 【図面の簡単な説明】

【図1】本発明の音声符号化方法における音源ベクトル生成手段の一実施例の構成を示すブロック図

【図2】本発明の音声符号化方法におけるピッチ周期の分析回数を示す模式図

【図3】従来の音声符号化方法におけるピッチ音源ベクトル生成手段の構成を示すブロック図

#### 【符号の説明】

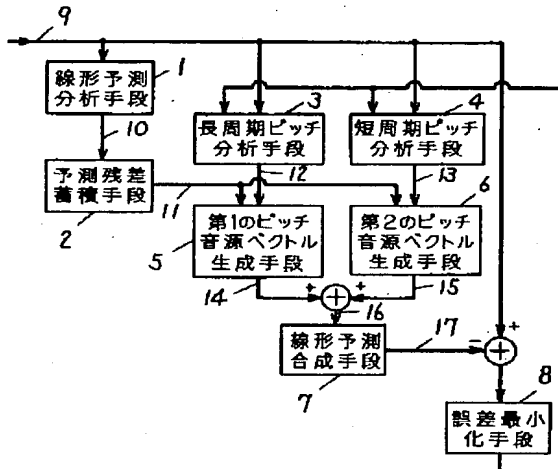
- 1 線形予測分析手段
- 2 予測残差蓄積手段
- 3 長周期ピッチ分析手段
- 4 短周期ピッチ分析手段
- 5 第1のピッチ音源ベクトル生成手段
- 6 第2のピッチ音源ベクトル生成手段
- 7 線形予測合成手段
- 8 誤差最小化手段
- 9 入力音声
- 10 予測残差

- 1 1 遅れ予測残差
- 1 2 長周期ピッチ
- 1 3 短周期ピッチ
- 1 4 長周期ピッチ音源ベクトル

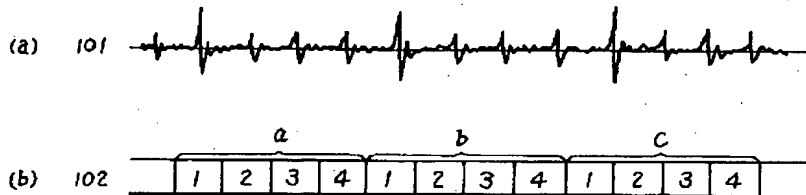
- 1 5 短周期ピッチ音源ベクトル
- 1 6 ピッチ音源ベクトル
- 1 7 合成音声

【図1】

- 3 長周期ピッチ分析手段
- 4 短周期ピッチ分析手段
- 9 入力音声
- 10 予測残差
- 11 遅れ予測残差
- 12 長周期ピッチ
- 13 短周期ピッチ
- 14 長周期ピッチ音源ベクトル
- 15 短周期ピッチ音源ベクトル
- 16 ピッチ音源ベクトル
- 17 合成音声



【図2】



【図3】

- 307 入力音声
- 308 予測残差
- 309 遅れ予測残差
- 310 ピッチ周期
- 311 ピッチ音源ベクトル
- 312 合成音声

